ACER

Australian Council for Educational Research

LSAY longitudinal surveys of Australian youth

SAMPLES AND SAMPLING FOR THE Y95 LSAY COHORT

Mike Long

Technical Paper Number 8

Longitudinal Surveys of Australian Youth

The 1995 Sample

April 1996

This paper outlines the intended and achieved samples for the first cohort in the program of *Longitudinal Surveys of Australian Youth*. Some earlier discussions that relate to the sample design are available in LSAY Technical Reports Numbers 2 and 7.

The sample design

The design was intended to provide a national stratified sample of Year 9 students which would permit a sample of some 10,000 young people to be interviewed by phone in late 1997. The major stratum considered in the design was State of schooling. Students from smaller states were to be over-sampled and, correspondingly, students from larger states were under-sampled. Selection of students within States was to be proportional by Sector. Three sectors were used as strata: Government schools, Catholic schools and Non-government, Non-Catholic (referred to as *independent*) schools. The population data for strata were taken from the *Schools Australia* series (ABS). Within strata, schools were to be selected proportional to their size. Information on the number of Year 9 students in each school came from ACER's Sampling Frame which, in turn, was based on information provided by the relevant State authorities and, in the case of non-government schools, by DEET. These figures were from the 1994 annual school census. Within schools two classes were to be randomly selected (again, proportional to their size). Schools were asked for a list of the number of students enrolled in each of their Year 9 classes for a subject studied by all Year 9 students in the school (usually English classes). Responses would be weighted to correct for the disproportionate sampling between strata and to correct for the variation between strata due to differential response rates and variable class sizes.

There was some uncertainty about class sizes, but it was assumed that two classes would yield a minimum of 35 students per school, which in turn would be consistent with an overall sample of some 10,000 students. It was anticipated that slightly larger class sizes would lead to a sample of around 12,000 students. Given sample attrition (including refusal and inability to contact identified in a 1996 mail survey), a sample of around 10,000 young people could be identified for phone contact in 1997.

An additional sample of some 500 Year 9 students was to be selected by sampling a further class from some schools. This group was to form the basis for a pilot sample in subsequent contacts with the cohort.

The sampling process

Lists of schools for each sector within each State were sorted by postcode. The cumulative total of Year 9 students was calculated. The interval required to yield the designed number of schools was determined, a random start made within that interval, and then the interval was applied to

select schools from the list. This process yields selection proportional to the size of the school and an implicit stratification by geography because of the postcode-order of the list. If constant numbers are selected from each school, the sample within each State becomes self-weighting.

Schools with less than 30 Year 9 students were combined with other small schools to form a pseudo-school with more than 35 Year 9 students. The selection of one school within a pseudo-school implied the selection of other schools within the pseudo school. Without the formation of pseudo-schools, there was the likelihood that the achieved sample would be less than the designed sample because 35 Year 9 students would not be able to be selected from smaller schools.

In practice, the procedure for selection of schools was modified from that outlined above. The *Third International Mathematics and Science Study* (TIMSS) -- another ACER project -- had already approached some schools in 1995. These schools were removed from the list before selection began. This had a relatively small effect on sample selection. This study focused on schools with Year 12 students. Selection in states with senior colleges (mainly Tasmania and the Australian Capital Territory) was therefore relatively unaffected and few schools were involved in the other small States. In the larger States, proportionately fewer schools were involved.

The sample for another ACER project, *Gender and School Education (GASE),* was drawn in conjunction with the LSAY sample in order to minimise instances in which schools were approached to participate in both studies and maximise response rates. The selection procedures were similar for the two studies. Essentially the interval of selection was altered so that for any given sector within State the number of schools required for both studies was selected and then allocated systematically to either LSAY or GASE. The details of this procedure differed between States.

Other procedures also differed between States. In New South Wales, Victoria and Queensland twice the required number of schools were selected. Schools were paired, so that in the event of refusal, the second school could be used as a replacement. A similar approach was used in the Australian Capital Territory, except that the replacement schools were those participating in the GASE study. In Western Australia a replacement sample of 50 per cent of the required number of schools was selected. In other States, there were too few schools to permit this approach to be used effectively. In the event of refusal, the next available school on the list was to be approached as the replacement.

The first contact with a school was by a letter to the principal. The letter was accompanied by a form and a reply-paid envelope. The form allowed schools to indicate whether or not they would participate in the project and, if they agreed, the class structure of their Year 9 students. Several weeks was allowed for the request to be considered by internal staff committees or school councils. If, after several weeks, no reply was received, a follow-up fax was sent to the principal. Again, if no response was received, a second fax was sent. In some cases a further series of phone calls was required. Frequently the first or second fax elicited a request that the initial letter be sent again.

It became clear that if an explicit refusal was required from a school before approaching a replacement, the sample would fall short of its target number of schools. In the New South Wales, Victoria, Queensland and the Australian Capital Territory government sectors, all replacements were approached. Follow-up procedures were employed until the required number of acceptances was achieved.

An industrial dispute in Western Australia between the teacher unions and the Ministry produced a very high initial refusal rate in the government (and possibly the Catholic) sector for that State. A ban on all non-teaching related activities meant that it was difficult for principals to elicit teacher support for the LSAY project. Schools were offered a $150 payment to employ a casual teacher to explicitly undertake the work required to organise LSAY at the school. This was not a universally accepted solution, but some principals and their teaching staff were prepared to allow their schools to participate in the project on this basis. In any event, all replacement schools were approached as were a further sample of 20 schools (half the schools that were not already included in LSAY or another ACER study).

There was also a low initial response rate from government and independent schools in South Australia. This was a particular problem for two reasons. First, there were relatively few replacement schools available in South Australia, in part because of the higher sampling fraction of schools in that State and in part because a number of schools already involved in projects with the South Australian Department of Education, Employment and Training were removed from the list at the request of the Department. Second, parents of students in government schools were required to give active consent to their child's participation. A higher refusal rate within schools was anticipated.

Two approaches were used to overcome these problems. First, a consultant was employed to contact schools in South Australia. Robert Slater, a retired South Australian primary school principal and past member of ACER Council proved effective in asking schools to reconsider an initial refusal or, if they had not responded, in agreeing to participate. Second, once schools agreed to participate, students were sampled at a higher rate than in other States.

Government schools in Tasmania also had a high initial refusal rate. Subsequent phone contact with principals indicated a concern about the number of requests for participation in research projects in addition to what they felt were high levels of reporting to the Ministry. There was an explicit desire to protect teaching programs from the intrusion of non-program related matters. In this context, the offer of money to employ a part-time teacher was not effective. The Department of Education in Tasmania provided a list of high schools which were not involved in any research projects and these schools were approached, again with a comparatively low acceptance rate.

Selection within schools proceeded as intended. In schools where selection of two classes did not yield 35 students, a third class was selected. In several schools, the principal requested that all students in Year 9 be tested. Table 1 outlines the designed and achieved samples.

Table 1 Designed and Preliminary Achieved Samples and Response Rates

State	Sector	Designed Schools	Schools Contacted	Achieved Schools	Response Rate	Designed Students	Achieved Students	Over-Sampling
NSW	***Govt***	46	92	44	47.8	1610	2118	31.6
	Cath	14	17	13	76.5	490	663	35.3
	Ind	6	7	6	85.7	210	309	47.1
	Total	66	116	63	54.3	2310	3090	33.8
VIC	***Govt***	38	76	39	51.3	1330	1931	45.2
	Cath	12	13	11	84.6	420	568	35.2
	Ind	8	10	8	80.0	280	366	30.7
	Total	58	99	58	58.6	2030	2865	41.1
QLD	***Govt***	34	68	37	54.4	1190	1809	52.0
	Cath	9	10	8	80.0	315	396	25.7
	Ind	7	14	7	50.0	245	319	30.2
	Total	50	92	52	56.5	1750	2524	44.2
SA	***Govt***	21	39	26	66.7	735	1050	42.9
	Cath	5	5	5	100.0	175	240	37.1
	Ind	4	6	5	83.3	140	430	207.0
	Total	30	50	36	72.0	1050	1720	63.8
WA	***Govt***	28	62	29	46.8	980	1187	21.1
	Cath	7	9	4	44.4	245	259	5.7
	Ind	5	6	6	100.0	175	391	123.0
	Total	40	77	39	50.6	1400	1837	31.2
TAS	***Govt***	15	28	11	39.3	525	337	-36.0
	Cath	3	3	3	100.0	105	136	29.5
	Ind	2	4	2	50.0	70	109	55.7
	Total	20	35	16	45.7	700	582	-16.9
NT	***Govt***	6	9	8	88.9	210	304	44.8
	Cath	1	1	1	100.0	35	50	42.9
	Ind	1	1	1	100.0	35	42	20.0
	Total	8	11	10	90.9	280	396	41.4
ACT	***Govt***	8	15	7	46.7	280	345	23.2
	Cath	4	8	4	50.0	140	205	46.4
	Ind	2	2	1	50.0	70	49	-30.0
	Total	14	25	12	48.0	490	599	22.2
AUS	***Govt***	196	389	201	51.7	6860	9081	32.4
	Cath	55	66	49	74.2	1925	2517	30.8
	Ind	35	50	36	72.0	1225	2015	64.5
	Total	286	505	286	56.6	10010	13613	36.0

Notes:

1 *Schools* refers to pseudo-schools.

2 *Response rate* is response rate for schools.

3 *Designed students* is based on 35 students per school.

4 The extensive over-sampling for independent schools in South Australia and Western Australia resulted from (large) schools requesting that all students in Year 9 be sampled.

5 Includes students to be assigned to the pilot sample.

Recommendations for the 1998 sample

The proposed process of contact with a school and then contact with a designated replacement school upon refusal was difficult to implement within the time available. It would have been more effective to have had an initial level of approaches that allowed for a proportion of refusals.

The values in Table 1 indicate the level of over-contact that may be required. The national response rate for government schools was 56.6 per cent. This suggests that government schools should be over-contacted by a little less than 100 per cent. This figure is probably too high. The response rates for New South Wales, Victoria and Queensland were achieved in the context of stopping follow-up procedures when the required number of schools had agreed. If follow-up procedures been implemented thoroughly, response rates of the order of 60 per cent would have been achieved for all three States. The industrial situation in Western Australia also contributed to a low response rate. These considerations (taken together with suggestions outlined below) indicate that an over-contact rate of about 60 per cent may be adequate. ie if 10 schools are required, 16 are approached, on the assumption of a 62.5 per cent response rate.

The Catholic and independent sectors had response rates of 74.2 and 72.0 per cent respectively. This suggests that over-contact of around 35 per cent might be satisfactory. In all cases, the resulting number of schools should be rounded up to the next whole number.

All schools in the sample were offered their choice of an ACER monograph and periodical in return for their participation in the project. Surprisingly, this created administrative difficulties for the school (Who should choose, did their library already have the books, and so on). Schools frequently failed to nominate the books they required and there were delays while this information was obtained from schools. The books also complicated the process of packing and dispatch at ACER. It would be simpler, and possibly have a more positive effect on response rates, if all schools were simply offered $100 in return for their participation in the project.

Table 2 shows the average class size for schools that provided this information and which were not part of a pseudo-school. This table provides one of the few sources of information about actual class sizes (rather than, say, eft enrolments divided by eft staff). In this context, the point of greatest interest is that within sectors within States, mean class sizes vary only marginally from about 25 students. Hence two classes should yield about 50 students. This means that in order to obtain a sample of say 12,000 students, it should only be necessary to sample some 240 schools (or pseudo-schools). This is a reduction from the 286 schools intended to be sampled in 1995.

There are other issues to consider. First, even though the average class size may be 25 students, the achieved sample within classes may be smaller due to refusal and absenteeism. This, however, may be more than compensated for by imposing a minimum sample size within each school (say 35 students) and selecting an additional class (or classes) if this is not achieved. Hence a lower bound is imposed on the variation in students per school and the average per school should be somewhat greater than 50 students. Second, the sample variation is dependent more on the number of schools than the number of students. Hence a reduction to 240 schools may be unwise, and a compromise number of (say) 250 more acceptable.

The extent of attrition between the selection of the sample and the first phone interview will be demonstrated for the 1995 sample by late 1997. This will provide a further guide on the extent to which allowance must be made for over-sampling in order to achieve an effective sample size of 10,000 for the start of the program of phone interviews.

Table 2: Average class size by state and sector

State	*Sector*	*Mean*	*Classes*	*Std. Dev.*	*Minimum*	*Maximum*
NSW	Govt	26.38	260	4.06	3	40
	Cath	25.43	67	5.40	10	35
	Ind	23.35	23	5.84	10	31
	All	26.00	350	4.53	3	40
VIC	Govt	24.04	267	2.23	13	27
	Cath	26.88	65	1.51	22	31
	Ind	25.69	35	2.75	22	31
	All	24.70	367	2.44	13	31
QLD	Govt	27.27	244	3.48	9	35
	Cath	26.72	40	3.67	17	32
	Ind	25.17	23	5.37	15	40
	All	27.04	307	3.70	9	40
SA	Govt	25.30	121	5.00	6	34
	Cath	27.26	19	3.11	23	31
	Ind	26.00	10	4.62	19	31
	All	25.59	150	4.80	6	34
WA	Govt	26.64	154	5.85	5	35
	Cath	28.65	23	6.50	10	34
	Ind	23.68	22	7.97	5	32
	All	26.55	199	6.27	5	35
TAS	Govt	24.65	40	4.91	7	29
	Cath	27.62	13	3.33	22	31
	Ind	29.00	6	0.00	29	29
	All	25.75	59	4.60	7	31
NT	Govt	23.50	24	4.93	17	33
	Cath	28.50	2	0.71	28	29
	Ind	23.75	4	2.22	21	26
	All	23.87	30	4.62	17	33
ACT	Govt	25.27	37	6.10	12	34
	Cath	26.42	26	5.81	12	33
	Ind	25.00	6	0.00	25	25
	All	25.68	69	5.70	12	34
AUS	Govt	25.79	1147	4.32	3	40
	Cath	26.67	255	4.39	10	35
	Ind	24.92	129	5.16	5	40
	All	25.86	1531	4.42	3	40

Notes

1 Includes only intact schools ie schools with an enrolment of 35 or more students.

2 Figures are based on estimates provided by principals and may involve some approximation.